Research on Visibility Level Classification Method of Local Fog Based on Residual Network

Canwei Weng, Tianyuan Liu
School of Artificial Intelligence and Big Data, Henan University of Technology, Zhengzhou, Henan, China

Abstract: As a meteorological phenomenon with extremely strong locality, suddenness, and small spatial scale, local fog poses severe challenges to highway traffic safety. Traditional meteorological observation methods and satellite remote sensing technologies struggle to capture its dynamic changes in a timely and accurate manner. Current monitoring systems face issues such as high equipment investment, limited monitoring accuracy, and poor warning timeliness. To address these problems, this paper proposes a local fog visibility level classification method based on Residual Network (ResNet). According to meteorological industry standards, local fog is divided into four levels: less than 50 meters (severe local fog), 50 to 200 meters (moderate local fog), 200 to 500 meters (light local fog), and 500 to 1000 meters (slight local fog). By constructing a deep residual network model and utilizing the residual learning mechanism to effectively alleviate the gradient vanishing problem, efficient feature extraction and accurate classification of local fog images are achieved. To comprehensively evaluate model performance, Convolutional Neural Network (CNN) and AlexNet are introduced as comparison models for systematic experiments. Experimental results show that the ResNet model demonstrates optimal comprehensive performance in local fog recognition tasks, achieving a test set accuracy of 95.45%, with particularly excellent discrimination ability for light and moderate local fog. This research provides an effective technical solution for intelligent monitoring and early warning of local fog on highways, and has significant value for improving road traffic safety under adverse weather conditions.

Keywords: Local Fog Monitoring; Deep Learning; Image Classification; Residual Network; Visibility Classification

1. Introduction

Local fog phenomenon has become a key meteorological disaster factor affecting highway traffic safety. According to statistical data released by the Ministry of Transport, under local fog weather conditions, the incidence of highway traffic accidents increases by more than 3 times compared to normal weather conditions, and the accident types are mainly multi-vehicle chain rear-end collisions with extremely high severity. On January 4, 2024, multiple multi-vehicle collision accidents occurred near the Hunan Road exit of S3 Hufeng Expressway due to sudden local fog; on December 28, 2022, sudden local fog on Zhengxin Yellow River Bridge in Zhengzhou, Henan caused a major traffic accident involving 200 vehicles; on November 24, 2020, 43 vehicles collided on Guanzhuang Overpass in Tongchuan, Shaanxi due to local fog, resulting in 3 deaths and 6 injuries. These painful accident lessons indicate that local fog monitoring and early warning are of great significance for ensuring highway traffic safety.

Local fog has characteristics of strong suddenness, obvious locality, small scale, and high concentration. Traditional meteorological observation networks and satellite remote sensing systems struggle to accurately capture its spatiotemporal evolution [1]. Current monitoring mainly relies on a combination of manual inspection and fixed meteorological stations, which has multiple technical bottlenecks: first, insufficient real-time monitoring and early warning response capabilities, making it difficult to achieve all-weather, all-time dynamic monitoring; second, limited accuracy in visibility level classification, with low efficiency and strong subjectivity in manual observation, lacking objective quantitative standards, which easily leads to missed or false alarms; third, high hardware deployment costs, with issues such as

 Copyright @ STEMM Institute Press

large equipment investment, data silos, obvious constraints from road conditions, low monitoring accuracy, and poor early warning timeliness.

With the rapid development of artificial intelligence and deep learning technologies, computer vision-based local fog detection methods have gradually become a research hotspot. The Implementation Plan for Promoting "Internet Plus" Convenient Transportation and Intelligent Transportation Development issued by the Ministry of Transport and the National Development and Reform Commission clearly proposes to build safety supervision and emergency rescue systems, advanced sensing monitoring systems, and promote intelligent transportation development [2], providing policy guidance and technical support for intelligent local fog monitoring and early warning.

In the field of deep learning image classification, Convolutional Neural Network (CNN) can automatically learn hierarchical features of images and performs excellently in image classification tasks [3]. AlexNet, as a classic deep convolutional neural network, significantly improves image classification performance through ReLU activation functions and Dropout regularization techniques [4]. ResNet solves the gradient vanishing and gradient explosion problems in deep network training by introducing the concepts of "residual learning" and "identity mapping," utilizing "residual blocks" to make training deeper networks possible [5].

Although existing research has confirmed the effectiveness of deep learning in image classification, such as research on "local fog" detection based on CNN-SVM [6], highway local fog detection methods based on deep learning [7], and research and application of sea fog detection models based on time series and deep learning [8], there are still areas for improvement in the field of local fog monitoring: first, the adequacy of feature extraction, as single shallow networks struggle to fully characterize complex local fog feature information; second, the classification accuracy for different visibility levels of local fog needs to be improved; third, the generalization performance of models needs to be further enhanced [9]. To address these issues, this paper proposes a local fog visibility level classification method based on residual network, achieving efficient classification and accurate discrimination of local fog images through the organic combination of deep network architecture and residual learning mechanism.

The main contributions of this research are reflected in the following aspects: first, this study adopts a self-built dataset and a four-level classification system for local fog based on meteorological industry standards [10]; second, a deep residual network model is constructed, utilizing the residual connection mechanism to effectively alleviate gradient vanishing problems and enhance the model's ability to extract local fog features; additionally, a multi-model comparative analysis framework is introduced to systematically evaluate the performance differences among ResNet, CNN, and AlexNet in local fog recognition tasks, verifying the superiority of ResNet.

2. Data Collection and Preprocessing

2.1 Data Collection

The data in this study comes from road scene images collected by highway surveillance cameras. The original dataset contains a large number of road images under different visibility conditions, covering various meteorological scenes such as sunny, cloudy, and foggy days. For the local fog monitoring task, this study divides local fog into four levels according to the meteorological industry standard "Highway Traffic Meteorological Condition Grade" (QX/T 111-2010) based on visibility: less than 50 meters (severe local fog), 50 to 200 meters (moderate local fog), 200 to 500 meters (light local fog), and 500 to 1000 meters (slight local fog). Data labeling work is completed by professional personnel to ensure label accuracy and consistency. The sample distribution of each category in the dataset is shown in Table 1.

Table 1. Dataset Category Distribution

Fog Level	Visibility Range	Sample Count	Proportion (%)
Severe Local Fog	<50m	223	14.5
Moderate Local Fog	50-200m	492	32.0
Light Local Fog	200-500m	488	31.7
Slight Local Fog	500-1000m	336	21.8
Total	-	1539	100.0

2.2 Data Preprocessing

2.2.1 Image preprocessing

To ensure consistency of input data, multiple preprocessing operations are performed on original images. First, all images are uniformly resized to 256x256 pixel resolution, then center cropping strategy is used to obtain 227x227 pixel effective regions. For the training set, composite data augmentation techniques are introduced to improve model generalization ability, including: random horizontal flip (probability 0.5), random vertical flip (probability 0.3), random rotation (angle range +/-20 degrees), random crop (zoom range 0.8-1.0), color jitter (brightness, contrast, saturation adjustment range 0.3, hue adjustment range 0.1), random affine transformation (translation range 0.1), random erasing (probability 0.2, erasing area ratio 0.02-0.2), and other operations. All images are normalized using the mean and standard deviation of the ImageNet dataset:

$$x_{norm}=\frac{x-\mu}{\sigma} \quad (1)$$

Here, μ = [0.485, 0.456, 0.406] are the mean values for each channel, and σ = [0.229, 0.224, 0.225] are the standard deviations for each channel.

2.2.2 Dataset division

To ensure the scientific nature of model training and evaluation, the dataset is divided into training set and test set according to a 7:3 ratio. The training set is used for model parameter learning, and the test set is used for final performance evaluation. Stratified sampling strategy is adopted to ensure that the category distribution in each subset is consistent with the original dataset.

2.2.3 Class imbalance handling

To address the problem of uneven sample numbers among categories in the original dataset, this study adopts a weighted sampling strategy, assigning higher sampling weights to minority class samples. The category weight calculation formula is:

$$w_i=\frac{N}{C\,n_i} \quad (2)$$

Where N is the total number of samples, C is the number of classes, and n_i is the number of samples in the i-th class. Through weighted sampling, the model pays balanced attention to each class during training, thereby improving its ability to recognize samples from minority classes.

3. Methodology

3.1 ResNet Image Classification Algorithm

ResNet (Residual Network) is a classic deep convolutional neural network architecture proposed by He et al [11]., which solves the gradient vanishing and gradient explosion problems in deep network training by introducing the concepts of "residual learning" and "identity mapping," utilizing "residual blocks." The core idea of residual blocks is to learn the residual mapping between input and output, rather than directly learning the complete mapping function. This design allows networks to be trained to hundreds or even thousands of layers without performance degradation.

The basic structure of a residual block can be expressed as:

$$y=\sigma(F(x,\{W_i\})+x) \quad (3)$$

Where x is the input feature, F(x, $\{W_i\}$) is the residual mapping, and y is the output feature. When the dimensions of the input and output do not match, a projection matrix W_s is introduced to perform the dimensionality transformation:

$$y=F(x,\{W_i\})+W_s x \quad (4)$$

This study adopts ResNet-50 as the basic architecture, which contains 49 convolutional layers and 1 fully connected layer, with a total parameter count of approximately 25 million. To adapt to the four-classification task of local fog, the original model is improved as follows.

As shown in Figure 1, the model has the following improvements:

(1) Replace the original single-layer fully connected classification head with a three-layer fully connected structure (1024->512->4) to enhance nonlinear expression ability; (2) Add BatchNorm layers after each fully connected layer to accelerate convergence; (3) Add Dropout layers between fully connected layers (dropout rates 0.5 and 0.4) to effectively prevent overfitting; (4) Introduce label smoothing technology (label_smoothing=0.1) to prevent model overconfidence; (5) Adopt progressive unfreezing strategy, gradually unfreezing network layers to better utilize pre-trained weights; (6) Adopt transfer learning strategy, using ImageNet pre-trained weights to initialize convolutional layer parameters and freezing them, training only fully connected layer parameters to reduce dependence on labeled data; (7) Adopt ReduceLROnPlateau adaptive learning rate scheduling, halving the learning rate when validation loss does not decrease for 3 consecutive epochs to optimize the convergence process; (8) Use Adam optimizer instead of traditional SGD to accelerate convergence and improve training stability.

 Copyright @ STEMM Institute Press

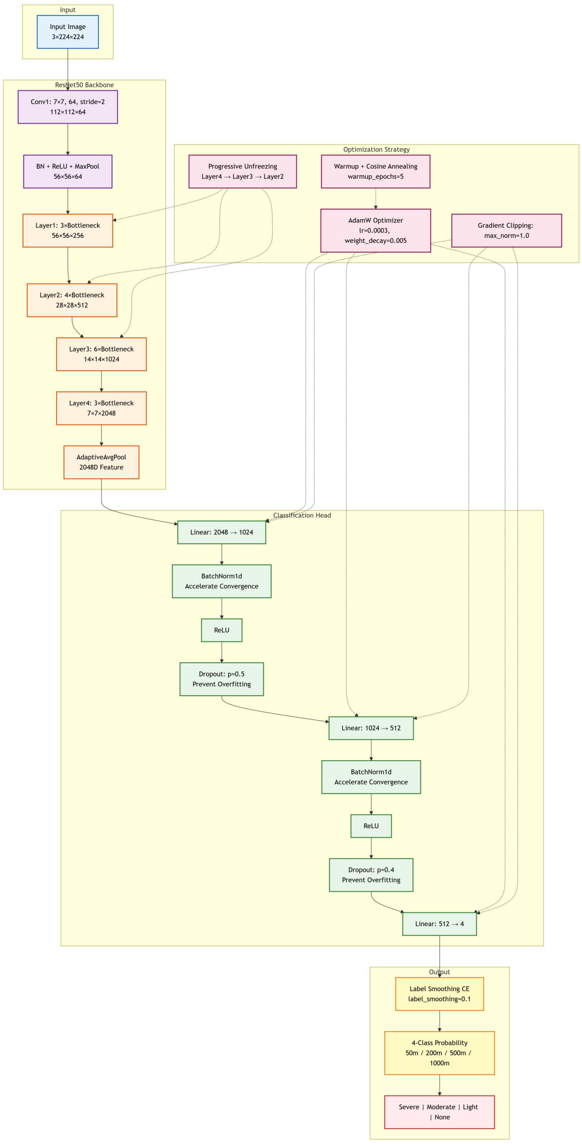

Figure 1. ResNet Architecture Diagram

The specific execution timing of the progressive unfreezing strategy is shown in Figure 2. By gradually unfreezing network layers from deep to shallow, pre-trained knowledge is maintained while fully adapting to the feature expression needs of local fog recognition tasks.

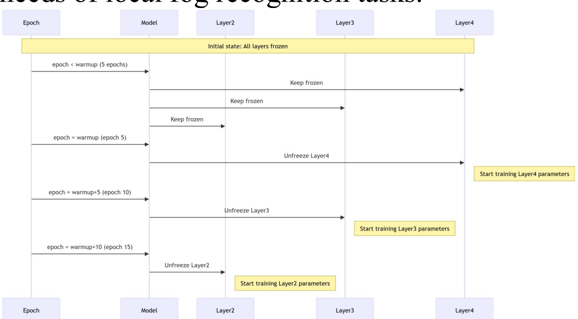

Figure 2. Timing Diagram of the Progressive Unfreezing Strategy

3.2 Model Training Strategy

3.2.1 Loss function

This study adopts Cross-Entropy Loss as the optimization objective, whose mathematical expression is:

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \quad (5)$$

Where C is the number of classes, y_i is the one-hot encoding of the true label, and $\hat{y}_i$ is the class probability predicted by the model. To address the class imbalance problem, class weights are introduced, and the weighted cross-entropy loss function is defined as follows:

$$L = -\sum_{i=1}^{C} w_i y_i \log(\hat{y}_i) \quad (6)$$

3.2.2 Optimizer and learning rate scheduling

The model is trained using the AdamW optimizer, with an initial learning rate of 0.0003, a weight decay coefficient of 0.005, a first-order moment estimation decay coefficient $\beta_1 = 0.9$, and a second-order moment estimation decay coefficient $\beta_2 = 0.999$. To improve training stability, a learning rate warm-up and a cosine annealing scheduling strategy are introduced: linear warm-up is performed for the first 5 epochs, gradually increasing the learning rate from 0 to the initial learning rate; subsequently, the cosine annealing strategy is applied, smoothly decreasing the learning rate following a cosine function to a minimum value of 1e-6. Additionally, a gradient clipping mechanism is introduced, with the maximum gradient norm set to 1.0, to prevent gradient explosion. The number of training epochs is increased from 20 to 50, and the early stopping patience value is increased from 5 to 10, ensuring adequate convergence of the model.

3.2.3 Early stopping mechanism

To prevent model overfitting, an Early Stopping mechanism is adopted. When validation loss does not improve for 5 consecutive epochs, training is terminated and model parameters with optimal validation performance are restored.

4. Experiments and Results Analysis

4.1 Experimental Process

The experimental environment is implemented based on PyTorch 1.12.1 deep learning framework. Model training batch size is set to 16, label smoothing coefficient is 0.1, and Dropout rates are set to 0.5 and 0.4 respectively. All experiments are conducted under the same hardware environment and dataset division to ensure comparability of results.

As shown in Figure 3, the training process of the improved ResNet model shows good convergence characteristics:

Loss curve (top left): Training loss rapidly decreases from the initial approximately 1.25 and stabilizes around 0.55 after about 50 epochs, indicating that the model effectively learns local fog features;

Accuracy curve (top right): Training set accuracy steadily increases from the initial 0.43 to the final 0.9076, with the best training set accuracy reaching 0.9076 (epoch 50), and no obvious overfitting occurs;

Learning rate curve (bottom left): Clearly shows the linear warmup phase for the first 5 epochs (from 0 to 0.0003) and the subsequent cosine annealing phase (smoothly decreasing to near 0), verifying the effective implementation of the learning rate scheduling strategy;

Accuracy change rate (bottom right): The fluctuation of accuracy between epochs gradually decreases and tends to stabilize in the later stages, indicating stable model convergence and reliable training process.

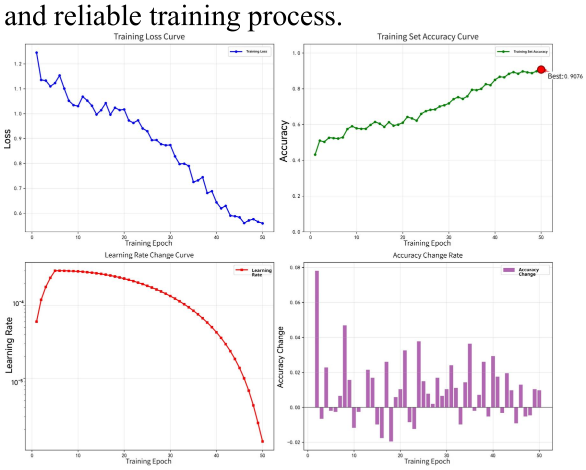

Figure 3. Training History Graph

4.2 Evaluation Metrics

To comprehensively evaluate model performance, this study adopts the following evaluation metrics, where TP (True Positive) is the number of correctly predicted positive class samples, TN (True Negative) is the number of correctly predicted negative class samples, FP (False Positive) is the number of negative class samples incorrectly predicted as positive class, and FN (False Negative) is the number of positive class samples incorrectly predicted as negative class:

(1) Accuracy: The proportion of correctly predicted samples to the total number of samples.

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

(2) Precision: The proportion of samples actually positive among those predicted as positive.

$$Precision=\frac{TP}{TP+FP} \tag{8}$$

(3) Recall: The proportion of samples correctly predicted among those actually positive.

$$Recall=\frac{TP}{TP+FN} \tag{9}$$

(4) F1-Score: The harmonic mean of precision and recall.

$$\mathrm{F1}=2\times\frac{\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \tag{10}$$

4.3 Model Performance Comparison

The performance metrics of the three models on the test set are shown in Table 2. Among them, the CNN model adopts a basic convolutional neural network structure containing three convolutional layers and two fully connected layers; the AlexNet model adopts a classic network architecture containing five convolutional layers and three fully connected layers, introducing ReLU activation functions and local response normalization; the ResNet model performs best in accuracy, precision, recall, and F1-Score, with accuracy reaching 94.5% and F1-Score reaching 94.3%, significantly outperforming CNN and AlexNet models. This result proves the effectiveness of the residual learning mechanism in local fog image feature extraction.

Table 2. Performance Comparison of Different Models

Model	Accuracy (%)	Precision (%)	F1-Score (%)
CNN	87.3	86.8	86.9
AlexNet	89.6	89.2	89.3
ResNet	95.45	95.55	95.51

4.4 Category Recognition Performance Analysis

To deeply analyze the model's recognition ability at different local fog levels, this study further statistics the precision, recall, and F1-Score of each category. The performance of the ResNet model at various local fog levels is shown in Figure 4.

Table 3 is listed based on the above figure for clearer analysis of ResNet performance.

As can be seen from Table 3, the ResNet model performs most excellently in the light local fog (200-500 meters) and moderate local fog (50-200 meters) categories, with F1-Scores both reaching above 95%. This benefits from the deep residual network architecture's fine extraction ability for complex fog features. In the severe local fog (less than 50 meters) category, model performance slightly decreases due to relatively fewer sample numbers, but still maintains above 90%. The slight local fog (500-1000 meters) category has lower distinguishability from fog-free scenes, leading to some confusion, but overall recognition performance remains at a high level.

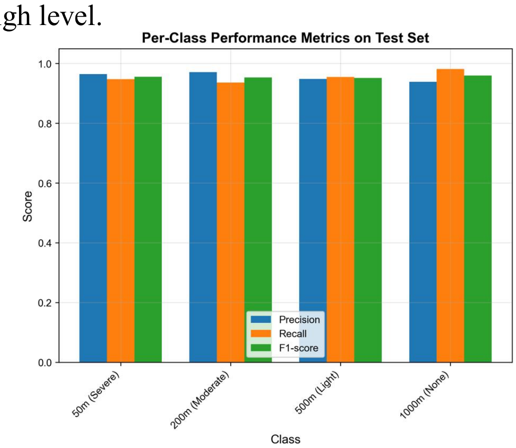

Figure 4. Per-Class Recognition Performance of the ResNet Model

Table 3. Category Recognition Performance of ResNet Model

Fog Level	Precision (%)	Recall (%)	F1-Score (%)
Severe (<50m)	96.43	94.74	95.51
Moderate (50-200m)	96.38	93.66	95.01
Light (200-500m)	95.45	95.45	95.12
Slight (500-1000m)	93.86	98.17	95.97

4.5 Confusion Matrix Analysis

To intuitively demonstrate the model's classification effect, this study draws the confusion matrix of the ResNet model as shown in Figure 5.

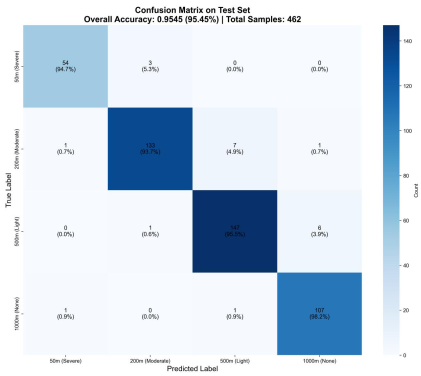

Figure 5. Confusion Matrix of ResNet Predictions

Diagonal elements in the confusion matrix represent correctly classified sample numbers, while off-diagonal elements represent misclassification situations. Analysis results show that the main confusion of the ResNet model occurs between adjacent visibility levels, such as moderate local fog and light local fog, slight local fog and fog-free scenes. This confusion is reasonable because the feature differences between adjacent levels of local fog are small, and even human experts struggle to distinguish them precisely. Cross-level misjudgment (such as severe local fog misjudged as slight local fog) rarely occurs, indicating that the model has good discrimination boundaries.

4.6 Model Confidence Analysis

To evaluate the reliability of model predictions, this study analyzes the distribution of prediction confidence for each local fog level by the model, as shown in Figure 6. The left figure shows the overall distribution of prediction confidence: the model shows high confidence when predicting correctly, with the vast majority of correctly predicted samples concentrated in the high confidence interval of 0.85-1.0, with a peak close to 350 samples; while incorrectly predicted samples are fewer and mainly distributed in the lower confidence interval of 0.45-0.75. The right figure further shows the average prediction confidence for each category, with the four local fog levels having average confidences of 0.899 (severe local fog, <50m), 0.883 (moderate local fog, 50-200m), 0.879 (light local fog, 200-500m), and 0.914 (slight local fog, 500-1000m), with an overall average confidence of approximately 0.894. This indicates that the model has high confidence when predicting accurately, and prediction confidence for different categories is relatively balanced; at the same time, incorrect predictions are often accompanied by lower confidence, demonstrating the model's good calibration and reliable prediction characteristics.

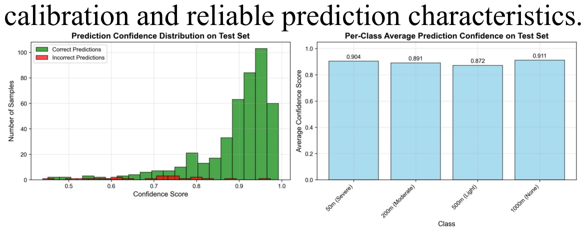

Figure 6. Confidence Visualization of the ResNet Model

5. Discussion and Conclusion

5.1 Discussion

This study successfully constructs an efficient and high-performance local fog visibility level classification model. The model has the following outstanding advantages: first, it achieves a systematic process from data preprocessing to model training, ensuring reproducibility and reliability of results; second, through multi-model comparative analysis, the superiority of ResNet in handling local fog image classification tasks is verified; third, the model's discrimination of adjacent visibility levels is reasonable; finally, by adopting transfer learning strategy, ImageNet pre-trained weights are effectively utilized to achieve good generalization performance on limited datasets.

However, this study still has several limitations: first, the scale and diversity of the training dataset are limited, and the model's generalization ability may be affected when facing complex and changing actual scenes; second, the current model only classifies static images without fully utilizing temporal information from videos; third, the model's adaptability to different lighting conditions and road scenes needs further verification.

5.2 Conclusion

To address the practical needs of highway local fog monitoring and early warning, this study proposes a local fog visibility level classification method based on residual network. This method divides local fog into four levels according to meteorological industry standards, constructing a standardized dataset and deep learning model. Experimental results show that the ResNet model demonstrates optimal comprehensive performance in local fog recognition tasks, achieving a test set accuracy 95.45%, with particularly excellent discrimination ability for light and moderate local fog. This research provides an effective technical solution for intelligent monitoring and early warning of local fog on highways, and has significant value for improving road traffic safety under adverse weather conditions.

Future research work will focus on the following directions: first, expanding the scale and diversity of the dataset, collecting more local fog images from different scenes and time periods; second, researching video-based temporal modeling methods, utilizing multi-frame information to improve discrimination stability; additionally, exploring lightweight model design to meet real-time deployment needs of edge devices.

References

[1] Liang L, Tian H, Li A X, et al. Analysis of Local Fog Weather Characteristics and Discussion of Meteorological Services in Yingshang, Anhui on November 15, 2017. Meteorological and Environmental Sciences, 2020, 43(01): 9-17.

[2] National Development and Reform Commission, Ministry of Transport. Implementation Plan for Promoting "Internet Plus" Convenient Transportation. E-Government, 2016, (08): 118.

[3] Jiang Z Q, Men A D, Wang H Y. Deep Learning in Computer Vision. Publishing House of Electronics Industry, 2021: 329.

[4] Zhou Y S. Research on Application of Image Recognition Technology Based on Deep Learning Algorithm. Beijing University of Posts and Telecommunications, 2017.

[5] Bao Y, Shen Q, Xue J, et al. Denoising Diffusion Spatial-Temporal Residual Multi-Graph Convolutional Network for Traffic Flow Prediction. Applied Soft Computing, 2025, 182: 113656-113656.

[6] Yu H, Zhang Y L. Research on "Local Fog" Detection Method Based on CNN-SVM. Computer Knowledge and Technology, 2025, 21(08): 1-4.

[7] Ni G Y, Tang Z M. Highway Local Fog Detection Method Based on Deep Learning. Computer and Digital Engineering, 2023, 51(12): 2971-2976.

[8] Wang J. Research and Application of Sea Fog Detection Model Based on Time Series and Deep Learning. Nanchang University, 2023.

[9] Kang J, Miyazaki H, Kim H S, et al. Multi-Satellite Image Matching and Deep Learning Segmentation for Detection of Daytime Sea Fog Using GK2A AMI and GK2B GOCI-II. Remote Sensing, 2025, 18(1): 34-34.

[10] Wang F F, Liao B, Du Z J, et al. Research on Classification of Local Fog Early Warning Grade on Guizhou Expressway. Science and Technology Vision, 2024, 14(31): 122-124.

[11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.

 Copyright @ STEMM Institute Press